A Simple Neural Network System for Wisconsin Card Sorting Test

Gülay B. Kaplan[†], İbrahim Genç[†], Cüneyt Güzeliş[‡]

gulay@btae.mam.gov.tr

[†]TÜBİTAK-MRC, ITRI, Gebze, Kocaeli, Turkey.

[‡]Dokuz Eylül University, İzmir, Turkey

Abstract—**In this paper, a model for mimicking the behavior of the prefrontal cortex in Wisconsin card sorting test is given for both healthy and lesioned subjects. A simple model based on winner take all network and multi layer perceptron suffices to model the affect of frontal lobe damage, which leads to perseveration as diminishing the influence of reinforcement.**

Keywords—**Prefrontal cortex, Wisconsin card sorting test, neural network, modelling**

I. Introduction

The prefrontal cortex (PFC) supports the cognitive functions that coordinate the execution of the most elaborate and novel actions of the organism. For this reason it has been called as the executive part of the brain and also the organ of creativity. It is the most interconnected region among all other parts of the brain. Thanks to these connections, PFC integrates inputs from sensory and associative areas of the cortex and information about internal mileu from limbic inputs [1]. So it is the highest stage of a hierarchy of neural structures dedicated to the representation and execution of the actions of the organism.

There is an extensive literature on differences in behavioral deficits from lesions in different subsections of the prefrontal cortex. The perseveration in behaviors that were formerly, but no longer rewarded, is one of the deficits causes from the damage of the dorsal area of the PFC [1]. The damages in this region diminish the influence of reinforcement on behaviour and perseveration is a characteristic action of such patients and best revealed by Wisconsin Card Sorting Test (WCST).

Since the PFC is responsible for the high level cognitive functions of the brain, it provides to engineers an inspiration for developing intelligent autonomous systems. In computational neuroscience research, scientists by employing the physiological findings on the PFC can develop some models which carry out some tasks of the PFC, and then these models may help to neuroscientist to improve their functional and systematical approach in their research.

There are some important modeling efforts in the literature help to understand how PFC works in WCST [2,3]. The one in [2] is based on a well known artificial neural network model namely ART and the other is a special neural network structure. This paper proposes a simple model which is based on multilayer perceptron cascaded by a winner take all network and models the PFC in WCST

for both normal and lesioned subject. The winner take all network is used as action network to select the cards. The multilayer perceptron defines a reinforcement type learning rule to adapt the decision criterion that the action network uses.

The paper is organized as follows; Section II and III gives information about the PFC and WCST, respectively. Section IV summarizes recent works on modeling WCST and finally, the proposed model is described in Section V.

II. The Prefrontal Cortex

The PFC which takes place at front of the motor and premotor cortices in the frontal lobes of the brain, encompass roughly 30% of the cerebral cortex in humans [4]. It is proportionally much larger in humans than in other species. This ratio is 3.5% in cats, and 17% in chimpanzees. The prefrontal region of the cortex is the most highly interconnected among all cortical regions. It receives affarent fibers from the brain stem, the hypothalamus, the limbic system, the thalamus and other areas from association cortices. It integrates inputs from sensory and associative areas of the cortex with inputs from motivational areas of the limbic system [1].

The PFC is the area of the brain most often associated with the executive processes in humans. The term "executive functioning" generally refers to the mechanisms by which the performance is optimized in situations requiring the operation of a number of cognitive processes [5]. Executive functioning is required when new effective plans of actions have to be formulated, and appropriate sequences of responses must be selected and scheduled. Components may include the enhancement of information held temporarily or 'on line' (Goldman-Rakic's concept of working memory [4]), the marshalling of attentional resources [6], the inhibition of inappropriate responses in certain circumstances [7] and the "monitoring" of behaviour with respect to affective or motivational state [1].

III. Wisconsin Card Sorting Test

In clinical diagnosis of prefrontal dysfunctions, some neuropsychological tests are used. Standardized clinical assessment of executive functions have traditionally involved tests such as the WCST, Stroop Test and Tower of Hanoi Test. The PFC has the strongest reciprocal connections with subcortical parts that are involved in processing in-

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle	Contract Number
A Simple Neural Network System for Wisconsin Card Sorting Test	
	Grant Number
	Program Element Number

Author(s)	Project Number
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
TUBITAK-MRC, ITRI Gebze, Kocaeli, Turkey	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-26, 2001 held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.

Abstract

Subject Terms

Report Classification	Classification of this page
unclassified	unclassified

Classification of Abstract	Limitation of Abstract
unclassified	UU

Number of Pages
4

ternal drive levels (the hypothalamus) and rewarding or punishing values of stimuli (the limbic system). For this reason, frontal lobe damages lead to diminished influence of reinforcement on behavioral performance. An aspect of this syndrome is perseveration in behaviors that were formerly, but no longer, rewarded. An example of perseveration occurs in the WCST.

In this test, there is a deck of cards and the subject is required to find the feature which distinguishes them. The cards are differentiated with respect to their three features namely color, shape, and number. Each feature can take four different values (color: red, green, blue, yellow; shape: triangle, star, cross, circle; number: 1,2,3,4). So there are 64 (4^3) different cards. The certain four cards are called as reference cards and the rest are called as response cards. The reference cards are distinguished from each other for all three features as shown in Fig 1. During the test, the reference cards are placed in front of the subject, and the experimenter takes a response card from the deck and presents to the subject. The subject puts this card on to the line of a reference card which has the same feature according to his/her own criterion. The experimenter gives a response as "right" or "wrong" without mention the correct category. The subject aims to get "right" response as soon as possible. Through out the test, the experimenter change his/her own criterion after a predetermined number of "right" responses of the subject.

Two common forms of the WCST are employed by neuropsychologists. The original test was introduced by Milner in which 2 decks of 64 cards are used and the criterion changed after 10 correct responses. The test is finished when the subject notices 6 criterion changes or 128 cards are used.
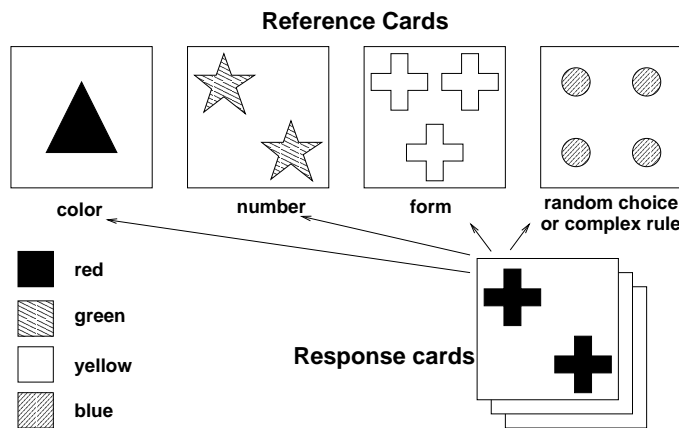
is following, and for the subject, who cannot determine for which rule he/she was enforced. Nelson eliminated those 40 cards and used only the 24 cards that shared only one attribute with each of the reference cards. Each card is used twice, for a total of 48 cards in response cards. *ii*) The experimenter changes criterion after 6 successful responses of the subject . *iii*) The subject is warned before each change of rules. *iv*) Finally, the correct rules are not defined a priori by the experimenter. Rather, for the first trial, the patient's criterion accepted as correct and also the second and the third criteria are defined by the experimenter so that all three criteria are different from each other. Then, these three criteria are employed periodically in the same order.

IV. NEURAL NETWORK MODELS FOR WCST

Leven and Levine simulated the card sorting data using the network of Figure 2 [2]. In this network based on adaptive resonance theory (ART), the nodes in the layer F1 code features, whereas the nodes in the F2 code categories. F1 divides input cards naturally into three subfields that code colors, numbers, and shapes. To each subfield corresponds a "habit node" and a "bias node" in addition to the F1 and F2. The habit nodes detect how often classifications have been made rightly or wrongly, on the basis of the given criterion. The bias nodes are affected both by habit node activities and by reinforcement signals. Reinforcement is given by the experimenter in the form of a statement of "right" or "wrong". A network parameter which corresponds to the gain of reinforcement signals to bias nodes was varied. Network with high gain acted like normal subjects, whereas the network with low gain acted like frontal patients.
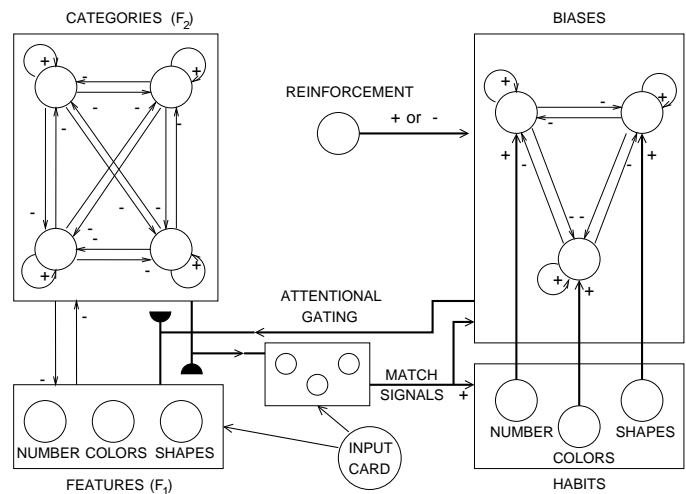

Fig. 1. Material used in the WCST.

Fig. 2. Leven and Levine's model for WCST.

Nelson introduced a number of modifications to the test [8]: *i*) Among the 64 cards of complete deck, 40 cards are ambiguous because two or more of their attributes are shared with the same reference card. For example, the card with two red triangles is similar to the first reference card in both color and shape. Such cards are ambiguous both for the experimenter, who cannot infer which rule the subject

Besides the model of WCST introduced Leven and Levine, the other major neural network model of this task was developed by Dehaene and Changeux [3]. Figure 3 shows the Dehaene-Changeux model. The analogies between the "input" and "memory" nodes of Figure 3 and "feature" and "category" fields of Figure 2, and rule coding

clusters and bias node are extremely close. In both models, the activity of rule clusters is an analog variable that can increase or decrease from one card presentation to the next. Dehaene and Changaux added to their model a feature they called "episodic memory" though it differs somewhat from the common usage of that term by psychologist. Their version of episodic memory kept track of rules that had been previously tried and not led to reinforcement, selectively reduced the activation of nodes representing such rules.
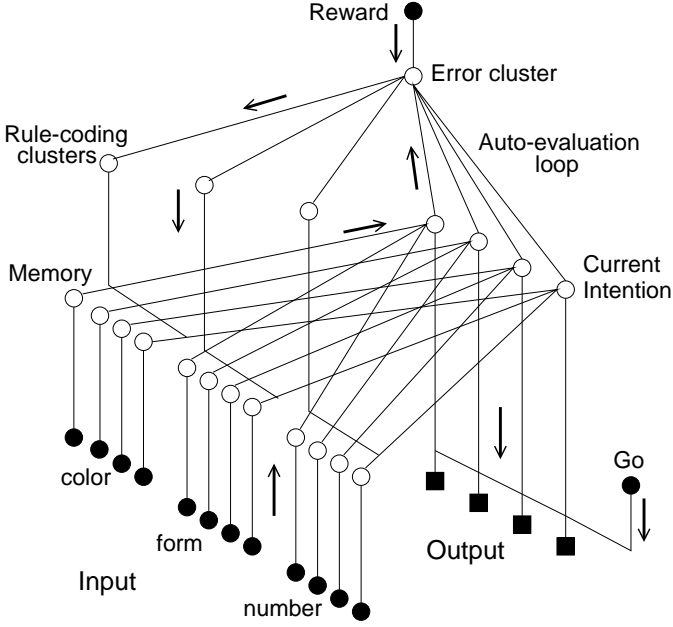
Features					
Color		Shape		Number	
Red	1 0 0 0	Triangle	1 0 0 0	1	1 0 0 0
Green	0 1 0 0	Star	0 1 0 0	2	0 1 0 0
Yellow	0 0 1 0	Cross	0 0 1 0	3	0 0 1 0
Blue	0 0 0 1	Circle	0 0 0 1	4	0 0 0 1

Fig. 3. Schematic architecture of Dehane-Changeux neural network.

Fig. 4. Proposed system for modeling WCST.

V. PROPOSED NEURAL NETWORK MODEL FOR WCST

A mixed procedure of Milner's and Nelson's versions of WCST is used in the model such that the 24 cards which sharing only one attribute with each of the reference cards are used and the experimenter's criterion is changed after 6 successful response of the subject. Moreover, subject is not warned before each change in the experimenter's criterion. The experimenter begins the the test with color criterion and changes it after each successful 6 responses of the subject in the order of color, shape and number repeatedly.

In the model the cards are represented as binary matrix of 12 (4 × 3) elements. Since the cards discriminated by their three features (color, shape and number) and the each feature could have four different values, each feature is coded in 4 digits and they are listed in 3 rows. Corresponding codes for each feature is depicted in Table 1.

The proposed model utilizes the subject behaviour by using two neural networks shown in Figure 4. The subject behaviour realized in two steps, one is decision and the second one is action. The decision network is designed to determine the selection criterion and the action network to select the reference card by using this criterion. The output of the action network shows a reference card which
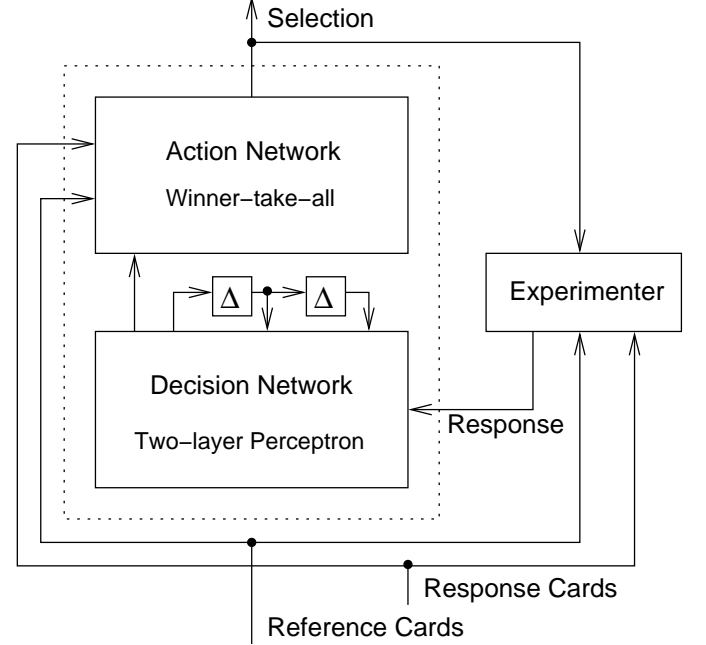
the subject has chosen. The action network, is a winner-take-all network composed of 4 neurons corresponding to the reference cards and each neuron takes input vector with 12 elements formed by concatenating the each row of the response card matrices. The neurons compute a weighted norm of difference between input card and reference card, as given in equation 1.

$$\|\mathbf{X} - \mathbf{T}^i\|_w = \frac{1}{2} \sum_{k=1}^{3} w_k \cdot \left[\sum_{j=1}^{4} |x_{kj} - t_{kj}^i| \right] \qquad (1)$$

where $\mathbf{T^i}$ is used for ith reference card and $\mathbf{X}$ for response cards. The index of reference card with minimum distance is chosen for the output of the action network. $\mathbf{w}$ represents the selection criterion vector. As the cards are discriminated by using three feature, i.e., color, pattern and number, this vector can take three different values as shown below.

$$\mathbf{w} = [1\ 0\ 0]\ \text{for color,}$$
$$\mathbf{w} = [0\ 1\ 0]\ \text{for shape,}$$
$$\mathbf{w} = [0\ 0\ 1]\ \text{for number.}$$

The decision network's output gives the subject's selection criterion for the next card. In other words, it de-

cides the current criterion should be kept or changed and if should be changed, it also determines the new value of it. The decision network takes current and two past selection criteria at the time $t-1$ and $t-2$ because the subject should remember the past selections to determine a new criterion, and the experimenter's response as inputs. Thus inputs of the decision network constitutes 10-elements binary vector. In the hidden layer of used two layer perceptron, 4 neurons are employed. Obviously, output layer consists of 3 neurons to represent the selection criterion, **w**. Since the decision network determines the selection criterion, it should be trained for all possible selections and experimenter's response. In the training set, the output vectors are the desired selection criterion that corresponds to the input vectors. While generating the training set, it is assumed that the subject changes his/her criterion if the experimenter's response is 0, in the order of color, shape and number, without loosing generality.

In the test procedure, response cards are selected randomly. For the first step, the decision criterion of the subject is also randomly assigned and it is determined by the decision network for the rest of the test. A response card is entered to the system, the action network selects a reference card by using the selection criterion which is the output of the decision network. After forming the action network output, the experimenter's response is generated as 1 if the subject selection is correct (match with experimenter's criterion), 0, otherwise. The response of the experimenter sets decision network into action and the decision network decide whether should keep or change the selection criterion. The current criterion, the past two criteria and the experimenter's response form the input of the decision network. The decision network generates the output using weights which are determined in the training phase. After forming the selection criterion, a next response card is selected as a new input.

In the proposed model, after a successful training phase, a set of weights of the decision network is found for simulation of a healthy subject, whose responses listed in Table II.

To simulate the perseverated response of the prefrontal lesioned patient some weight values of the decision network are modified adaptively. The reason of perseveration might be the fact that the lesioned subject more emphasize on his past selections instead of the experimenter's response. Because the frontal lobe damage leads to diminished influence of reinforcement on behavioral performance. The perseverative results are obtained when the weights between hidden neurons and experimenter's response and the biases of the same neurons are modified. The modifications of the weights are in a direction to decrease the effect of the experimenter's response. Changes in biases have an affect of resistance to excitatory effect of experimenter's response. These modifications on weights and biases are initiated by the rewarding responses of the experimenter. After the modifications are completed, the subject's selections are shown in Table II.

TABLE II
Responses of the Patient

Current Selection			Exp'er Response	Healthy Case			Lesioned Case		
0	0	1	0	0	1	0	0	0	1
0	0	1	1	0	0	1	0	0	1
0	1	0	0	1	0	0	0	1	0
0	1	0	1	0	1	0	0	1	0
0	0	1	1	0	0	1	0	0	1
1	0	0	1	1	0	0	1	0	0
0	0	1	1	0	0	1	0	0	1
0	0	1	1	0	0	1	0	0	1
0	1	0	1	0	1	0	0	1	0
0	0	1	1	0	0	1	0	0	1
0	1	0	1	0	1	0	0	1	0
0	1	0	1	0	1	0	0	1	0
1	0	0	1	1	0	0	1	0	0
0	0	1	1	0	0	1	0	0	1
0	1	0	1	0	1	0	0	1	0
0	1	0	0	1	0	0	0	1	0
1	0	0	0	0	0	1	1	0	0
1	0	0	1	1	0	0	1	0	0
0	1	0	1	0	1	0	0	1	0
0	0	1	0	0	1	0	0	0	1

VI. Discussions and Conclusions

This paper gives a simple structure to model PFC in performing WCST for healthy and lesioned subjects. The main idea behind the model is to construct the subsystems i.e., the experimenter's response as limbic system, the decision network as PFC and the action network as motor cortex and effect of them on each other in healthy and lesioned cases.

References

[1] J.M. Fuster, "Cognitive functions of the frontal lobes", in *The Human Frontal Lobes*, B.L. Miller and J.L. Cumming, Eds., pp. 187–195. The Guilford Press, 1999.
[2] S. Leven and D. Levine, "Effects of reinforcement on knowledge retrieval and evaluation", *IEEE First ICNN*, vol. 2, pp. 269–279, 1987.
[3] S. Dehaene and J.P. Changeux, "The wisconsin card sorting test: Theoretical analysis and modeling in a neuronal network", *Cerebral Cortex*, vol. 1, pp. 62–79, 1991.
[4] P. Goldman-Rakic, "Circuitry of primate prefrontal cortex and regulation of behavior by represential memory", in *Handbook of Physiology:The Nervous System*, F. Plum, Ed., vol. 5, pp. 343–417. American Physiological Society, 1987.
[5] A. Baddaley, *Working Memory*, Clarendon Press, Oxford, 1986.
[6] T. Shallice, "Spesific impairments of planning", *Phil. Trans. Roy. Soc. London*, vol. B, no. 298, pp. 199–209, 1982.
[7] T. Shallice and P. Burges, "Supervisory control of action and thought selection", in *Attention; Selection, Awareness and Control*, A. Baddeley and L. Weiskrantz, Eds. Oxford: Clarendon Press, Oxford, 1993.
[8] H. Nelson, "A modified card sorting test sensitive of frontal lobe defects", *Cortex*, vol. 12, pp. 313–324, 1976.